A higher order time-dependent

van den Berg, I.P.

A higher order time-dependent De Moivre-Laplace theorem

SOM theme A:
Structure, Control and Organization of Primary Processes

I.P. van den Berg*

Abstract

Viewed at an appropriate scale, the binomial coefficients $\binom{n}{j} p^j (1-p)^{n-j}$ with fixed n and varying j are close to the Gaussian curve. If we let also vary n, the family of binomial coefficients becomes close to a surface, described by a Gaussian function of two variables $G(t,x)$. This two-dimensional central limit theorem appears to be valid to higher order: certain difference quotients of binomial coefficients are close to the corresponding partial derivatives of $G(t,x)$. To prove that this transition from the discrete to the continuous is justified we use the stroboscopy method of nonstandard analysis. The theorem is useful in derivations of the heat equation and the Black-Scholes equation, starting from binomial stochastic processes.

Keywords: binomial coefficients, Gaussian function, Central limit theorem, S-continuity, stroboscopy.

* University of Groningen, Department of Econometrics, P.O. Box 800, 9700 AV Groningen, The Netherlands

1. Introduction, the main theorem

1.1 The main theorem

The main theorem states that the family of binomial coefficients viewed at an appropriate scale is close to a surface, whose sections are normal density functions, and that the closeness is valid up to higher order. Before formulating the main theorem explicitly, we introduce some notation and give an informal presentation.
Consider first the binomial coefficients.

$$B(N,j) = \binom{N}{j} \frac{1}{2}^N$$

with fixed, large N. Let us make the following rescaling

$$\begin{aligned} y &= \frac{j - N/2}{\sqrt{N}/2} \\ \beta_N(y) &= B(N,j) \cdot \sqrt{N}/2 \end{aligned}$$

Then by the usual De Moivre-Laplace central limit theorem the function $\beta_N(y)$ is close to the standard normal density function

$$\mathcal{N}(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Now let us look at the binomial coefficients $B(\nu, j)$ dynamically, with varying ν, placed in the form of a Pascal triangle, and rescaled such that the large number N becomes the new unity. To be precise, we carry out the rescaling

$$t = \frac{\nu}{N} \tag{1.1}$$

$$x = \frac{j - \nu/2}{\sqrt{N}/2} \tag{1.2}$$

$$b(t,x) = B(\nu, j)\ \sqrt{N}/2 \tag{1.3}$$

(See also fig. 2.1).
Then, as it is expected because at time t the standard deviation of the probability distribution $b(t,.)$ equals $\sqrt{t}$ the function $b(t,x)$ is close to the time dependent Gaussian density function, related to Standard Brownian Motion

$$G(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) \tag{1.4}$$

Using nonstandard analysis (we use the axiomatic theory IST of Nelson, see [8] or [9] for an up to date presentation), we can make the following precise statement: if N is infinitely large, for all positive appreciable t and limited x of the form (1.1) and (1.2) we have the approximation

$$b(t,x) \simeq G(t,x)$$

What is more, such an approximation still holds up to higher order. Certain partial difference quotients of $b(t,x)$ appear to be infinitely close to certain partial derivatives of $G(t,x)$. In order to formulate this result, we introduce the notations

$$\begin{aligned}
\delta t &= 1/N \quad , \quad \delta(2t) = 2/N \\
\delta x &= 1/\sqrt{N} \\
\delta_1 b(t,x) &= b(t+2\delta t, x) - b(t,x) \\
\delta_2 b(t,x) &= b(t, x+\delta x) - b(t,x) \\
\delta^2_{2,2} b(t,x) &= b(t, x+2\delta x) - 2b(t, x+\delta x) + b(t,x).
\end{aligned}$$

If again N is infinitely large, we have for every appreciable t of the form (1.1) and every limited x of the form (1.2)

$$\begin{aligned}
\frac{\delta_1 b(t,x)}{\delta(2t)} &\simeq \frac{\partial G}{\partial t}(t,x) \\
\frac{\delta_2 b(t,x)}{\delta x} &\simeq \frac{\partial G}{\partial x}(t,x) \\
\frac{\delta^2_{2,2} b(t,x)}{(\delta x)^2} &\simeq \frac{\partial^2 G}{\partial x^2}(t,x).
\end{aligned}$$

The main result concerns binomial coefficients

$$B_p(N,j) = \binom{N}{j} p^j (1-p)^{N-j}$$

with p slightly different from $\frac{1}{2}$. Indeed, put

$$\begin{aligned}
p = p(a) &= \tfrac{1}{2} + \frac{a}{\sqrt{N}} = \tfrac{1}{2} + a.\sqrt{\delta t} \\
b_a(t,x) &= B_p(\nu, j) \cdot \sqrt{N}/2 \\
G_a(t,x) &= \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-2at)^2}{2t}\right)
\end{aligned}$$

Then we have

Theorem 1.1 *(Main theorem). Let N be infinitely large and $a \in \mathbb{R}$ be limited. Then we have for all appreciable t of the form* (1.1) *and limited x of the form* (1.2)

$$
\begin{aligned}
b_a(t,x) &\simeq G_a(t,x) && (1.5)\\
\frac{\delta_1 b_a(t,x)}{\delta(2t)} &\simeq \frac{\partial G_a}{\partial t}(t,x) && (1.6)\\
\frac{\delta_2 b_a(t,x)}{\delta x} &\simeq \frac{\partial G_a}{\partial x}(t,x) && (1.7)\\
\frac{\delta^2_{2,2} b_a(t,x)}{(\delta x)^2} &\simeq \frac{\partial^2 G_a(t,x)}{(\partial x)^2} && (1.8)
\end{aligned}
$$

To get an idea of the domain of validity of the main theorem we remark first that the result obviously does not hold for very small t. Indeed if $t = \nu/N$ with standard fixed ν we deal with standard binomial coefficients, certainly not infinitely close to a Gaussian curve. If t is appreciable and x is infinitely large the result is meaningless for $|\ x\ | \geq t/\sqrt{\delta t}$, for which $b_a(t,x)$ is undefined. For $|\ x\ | \leq t/\sqrt{\delta t}$ the result holds but contains little information stating just the near equality of two infinitesimals. However, with respect to theorem 1.1 a weaker result holds: all the above expressions are quadratically-exponentially small, i.e. of the form $c\exp(-bx^2)$ with c limited and b appreciable (see theorem 6.1). For unlimited t theorem 1.1 continues to hold but like for unlimited x it is almost empty, all quantities involved being infinitesimal.

1.2 Use of the main theorem

The results of this paper contain in a sense a derivation of the density function of standard Brownian motion by straightforward means: elementary asymptotics and stroboscopy, a nonstandard method of going from the discrete to the continuous. The latter method, though intuitive, is not elementary because it uses the notion of "shadow"; however it avoids the use of non-elementary limits and topology. The fact that the closeness between the family of binomial coefficients and the Gaussian surface is higher order is applied in the following way.

It is well known that to a binomial stochastic process we may associate in a natural way a partial difference equation, describing the discrete "surface", union of all possible trajectories (it is very simple to derive this equation, see for instance section 7, especially fig. 7.1 and formula (7.2)). For binomial stochastic processes with small steps this equation looks very much alike to a discrete heat equation (see also formula (7.3)). In fact the usual heat equation as a partial differential equation describes a

smooth surface close to the surface of discrete trajectories.
It appears that the extended De Moivre-Laplace theorem 1.1, together with the general results on the infinitesimal approximation of partial difference quotients and partial differential quotients of [4], enables us to make the transition from binomial stochastic processes to the heat equation in a straightforward way, by intermediary of the above mentioned partial difference equation. Thus we avoid either the complications of continuous-time stochastic processes, Fourier analysis, distributions and partial differential operators of classical rigorous derivations, [11], [12], or the ad-hoc and unproved assumptions on smoothness of the involved functions - of which one knows only a discrete sample - of the usual informal derivations.
We will apply the theorem in [2], where we derive the heat equation from binomial processes with rather arbitrary initial conditions. Furthermore, we will apply the theorem in the context of the discrete backward stochastic processes of option pricing theory, i.e. the well-known Cox-Ross-Rubinstein model from which we will derive directly the Black-Scholes partial differential equation [6]. In a more calculatory way the latter result is also obtained in [13]. This study, like [5], also contains a derivation of the Black-Scholes option price formula in this setting.

1.3 Outline of this paper

The mathematical context of our approach and preliminary results are presented in section 2. First (section 2.1) we define the binomial network to which the binomial coefficients are rescaled. Section 2.3 contains a brief exposition of the terminology of nonstandard asymptotics, and some preliminary approximation lemma's needed for the asymptotic estimations of the remaining sections. In section 2.4 we recall the nonstandard tools, based on the notions of S-continuity and stroboscopy, which make our approach to the transition from the discrete to the continuous possible. In section 3 we present estimates of the difference quotients of the binomial coefficients. Crucial are the asymptotic lemma's 3.1 and 3.2. Using these lemma's we obtain approximate difference equations for the binomial coefficients.

These difference equations will be solved using the stroboscopy methods of section 2.3. We obtain the time-dependent De Moivre-Laplace theorem in first approximation in section 4, and to higher order approximation in section 5, thus proving the main theorem.
In section 6 we show that these approximations do no longer hold for certain binomial coefficients with high arguments. However, they may be suitably weakened, to a form which is still useful.